DOCUMENT RESUME

ED 156 711 TM 007 197

AUTHOR Hakstian, A. Ralph; Atell, Robert A.
TITLE Some Further Notes on the Problem of Oblique Factor Transformation.
INSTITUTION Alberta Univ., Edmonton. Div. of Educational Research.
REPORT NO RIR-73-1
PUB DATE Feb 73
NOTE 40p.; For related document, see EJ 041 287

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Comparative Analysis; Correlation; Data Analysis; *Factor Analysis; Factor Structure; *Oblique Rotation; *Statistical Analysis; *Transformations (Mathematics)

ABSTRACT

Four prominent oblique transformation techniques--biquartimin, direct oblimin, promax, and the Harris-Kaiser procedure--are described and compared as to optimality of solution. Additionally, two newly-developed procedures--falling into the category designated as Case III by Harris and Kaiser (1964)--are presented and included in the comparisons. The techniques are compared in light of their ability to yield clear simple structures over many data sets--some constructed and some "real"--varying widely in terms of number of variables and factors, factorial complexity, and clarity of the hyperplanes. Results indicate that no single computing procedure can be counted on to be uniformly optimal for all kinds of data. Also, there appear to be complicated interactions between the specification parameter and various characteristics of the data. The two new procedures yielded encouraging results. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED156711

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Division of Educational Research

University of Alberta

RIR-73-1

SOME FURTHER NOTES ON THE PROBLEM OF OBLIQUE FACTOR TRANSFORMATION

A. Ralph Hakstian and Robert A. Abell

University of Alberta

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E.W. Romaniuk

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

February, 1973

The contents of this report are not to be cited without permission of the authors.

- TM002 192 -

ABSTRACT

Four prominent oblique transformation techniques--biquartimin, direct oblimin, promax, and the Harris-Kaiser procedure--are described and compared as to optimality of solution. Additionally, two newly-developed procedures--falling into the category designated as Case III by Harris and Kaiser [1964]--are presented and included in the comparisons. The techniques are compared in light of their ability to yield clear simple structures over many data sets--some constructed and some "real"--varying widely in terms of number of variables and factors, factorial complexity, and clarity of the hyperplanes. Results are discussed, and implications for practice are noted.

SOME FURTHER NOTES ON THE PROBLEM OF OBLIQUE FACTOR TRANSFORMATION

A. RALPH HAKSTIAN AND ROBERT A. ABELL

UNIVERSITY OF ALBERTA

1. Introduction

The research reported in this paper constitutes an extension of an earlier study by the senior author [Hakstian, 1971], the latter investigation being directed at comparatively evaluating the oblimax [Saunders, 1961], promax [Hendrickson & White, 1964], maxplane [Cattell & Muerle, 1960], and Harris-Kaiser [Harris & Kaiser, 1964] methods of oblique factor transformation. The large class of oblimin solutions [see Carroll, 1957; Harman, 1967; Jennrich & Sampson, 1966] was ignored in the earlier paper, and, thus, one objective of the present paper is to complete the comparison of prominent oblique transformation techniques by comparing solutions obtained by oblimin methods with those obtained by the promax and Harris-Kaiser methods, the latter two procedures constituting the most promising of those compared in the earlier study.

A second objective of the present research was to explore the possibility of establishing new principles of oblique transformation--based, perhaps, on the best features of the more promising existing techniques. To this end, two new techniques were developed, both constituting, to some degree, extensions of Harris and Kaiser's [1964] formulations. These extensions are described in the next section, and empirical results obtained with them are presented, in a later section, for comparison with those obtained by the established procedures examined. As in Hakstian's [1971] earlier study, the criterion of closeness of the obtained primary-factor

patterns to an ideal simple structure was employed in the present investigation. This criterion is operationalized in a later section.

A summary of earlier studies containing comparisons of analytic transformation methods is given in the paper by Hakstian [1971], and these studies are not reexamined here. Additionally, the reader is referred to the above paper for a discussion of methods of comparing the quality of several factor solutions. As before, several criteria were employed, in the present study, in assessing comparative quality among solutions, the most heavily weighted criterion again being closeness--in terms of factor angular separations--of an obtained set of factors to a set of "ideal" oblique axes, for the data at hand.

A description of the established oblique transformation procedures compared is given in Section 2, along with an outline of two new extensions that appear to show promise with empirical data. The methods of comparing solutions, and a description of the data matrices used, appear in Section 3, followed by the results of the comparisons and a discussion of these results in Section 4. Some conclusions and recommendations follow, in Section 5.

2. Procedures Compared

Four established methods of oblique transformation and two new modifications were compared in the present study. As is well known, the task in oblique factor transformation is to establish a matrix G, of order q factors x q, and, in general real and non-orthonormal, of unit-length columns of direction cosines, to be applied as a post-multiplier to some unrotated factor pattern matrix A, p variables x q factors. The term "factor pattern" is used here in a generic sense, as A may be a pattern matrix under the component, common-factor or image [Guttman, 1953] models.

The optimal G matrix, for the purposes of the present study, is that which transforms the orthogonal factors represented by A to as close an approximation as possible to a simple structure [Thurstone, 1947] configuration. This simple structure may be manifested in the primary-factor pattern matrix, P, p x q, given by

$$P = A(G')^{-1}. \tag{1}$$

Other matrices involving G are the primary-factor structure matrix, S, p x q, and the primary-factor intercorrelation matrix Φ, q x q:

$$\begin{aligned} S &= AG, \text{ and} \\ \Phi &= G'G. \end{aligned} \tag{2}$$

Several principles have been advanced over the years for the establishment of G; some of these are subsequently discussed.

Established Methods

The two techniques which, overall, tended to yield the best simple structures in the study by Hakstian [1971] were the promax and Harris-Kaiser methods, with the latter procedure generally the better of the two. In the promax procedure, A is first transformed to an orthogonal (usually varimax) simple structure, and then a "target" matrix is established from this preliminary transformed pattern by raising each element to some power b, and ensuring that the original algebraic sign is retained. The initial matrix A, is then fitted, in a least squares sense, to this target; the transformation is column-normalized, and a simple structure matrix in the reference system is obtained. Well-known column rescaling relationships permit the calculation of P in (1). Hendrickson and White [1964] recommended values of two or four for b--either value, of course, resulting in a target matrix in which the relative distance (in absolute value) between significant and insignificant loadings on a factor is enhanced. Hakstian [1971]

concluded that (1) the optimal value for b is a function of the factorial complexity of the preliminary orthogonal solution, with low complexity associated with larger values of b and vice versa, and (2) size of b is directly related to obliquity (departure from orthogonality) of the obtained factors; i.e., as b increases, so does the departure from orthogonality. In the present study, two solutions based on the promax rationale were obtained for each data set, those with b set to two and four, both based on preliminary normal varimax solutions.

The term "Harris-Kaiser procedure" was used in the earlier paper with reference to the practical suggestions resulting from Harris and Kaiser's [1964] generalized restatement of the oblique transformation problem. Specifically, the two special cases of the "Case II" paradigm--the (1) independent cluster solution and (2) $P'P$ proportional to Φ solution--were termed "Harris-Kaiser" solutions. Harris and Kaiser began by showing that for non-singular P matrices, equation (1) can be rewritten

$$P = QMT_1D_1T_2D_2, \tag{3}$$

where Q, $p \times q$, is a matrix of unit-length latent vectors, M, $q \times q$, is a diagonal matrix of square roots of latent roots, T_1 and T_2, $q \times q$, are real, orthonormal matrices, and D_1 and D_2, $q \times q$, are positive-definite diagonal matrices. Thus, if A is not in canonical (or principal axes) form, i.e., equal to QM, it can be rotated orthogonally to this form. The Case II formulation is obtained by setting T_1 to I, and rewriting MD_1 as D_3 so that (3) becomes

$$P = QD_3T_2D_2. \tag{4}$$

The matrices D_3 and T_2 are the unknowns in the Case II formulation, with D_2 a function of Q, D_3, and T_2. The critical unknown appears to be D_3, with T_2 being given by maximization of some orthogonal rotation function on

$\underline{Q}\underline{D}_3$. Equation (4) can be further specialized to

$$P = QM^{c}T_2D_2, \tag{5}$$

where $\underline{D}_3$ is seen as a function of $\underline{M}$. Empirical evidence suggests values of $\underline{c}$ between 0 and .75 as optimal, with the two special Case II formulations noted above given by $\underline{c} = 0$ (independent cluster solution) and $\underline{c} = .5$ ($\underline{P}'\underline{P}$ proportional to $\underline{\Phi}$ solution). Empirical evidence also suggests that the particular choice of orthogonal rotation criterion--used in generating $\underline{T}_2$--is less important, in terms of quality of obtained solution, that the value of $\underline{c}$. An interesting feature of the general Harris-Kaiser formulation, as represented by (3), is that by restricting $\underline{D}_1$ (and, as a result, $\underline{D}_2$) to the class of positive-definite diagonal matrices, transformation to singularity (factor collapse) is precluded. At a more practical level, this same feature is true for solutions represented by (5). The problem of transformation to singularity has plagued many earlier analytic oblique procedures. In the present study, the general Harris-Kaiser formulation was represented by two solutions for each data set: the (1) independent cluster and (2) $\underline{P}'\underline{P}$ proportional to $\underline{\Phi}$ versions, both utilizing the normal varimax criterion for generating $\underline{T}_2$.

Although parenthetical to the main purposes of this paper, the following digression may be useful, in connection with a discussion of the Harris-Kaiser method. Several kinds of factor solutions are obtained by diagonal rescaling procedures. Inverse uniqueness rescaling is used, for example, in obtaining canonical, maximum likelihood, and generalized least squares [Jöreskog & Goldberger, 1972] common-factor solutions. Inverse communality rescaling is employed in alpha factor analysis, and rescaling by the inverse anti-image variances (actually standard deviations) is employed in one variant of image analysis [Harris, 1962]. In such

solutions, the unrotated pattern matrix, $\underline{A}$, is given by $\underline{D_s WL}$, where $\underline{D_s}$, $\underline{p} \times \underline{p}$, is the inverse of the diagonal rescaling matrix and $\underline{W}$, $\underline{p} \times \underline{q}$, and $\underline{L}$, $\underline{q} \times \underline{q}$ and diagonal, are the matrices of, respectively, the $\underline{q}$ retained latent vectors and square roots of the $\underline{q}$ retained latent roots of the particular rescaled covariance matrix factored in each case. In the application of the Harris-Kaiser procedure to such factor solutions, $\underline{Q}$ in (5) is typically replaced by $\underline{W}$, $\underline{M}$, by $\underline{L}$, and the optimization of the orthogonal rotation criterion is carried out on $\underline{WL}^c$, with the scale factors, $\underline{D_s}$, applied after the orthogonal rotation has been performed. This practice results in a slightly less than optimal final solution, in that the basic referent for the rotation does not contain the influence of these scale factors, and the application of the latter may disturb, to some extent, the simple structure achieved by the rotation. A better basic referent for Harris-Kaiser solutions can be obtained by first orthogonally rotating $\underline{D_s WL}$ to canonical form $\underline{QM}$ by

$$D_s WL(D_s WL)' = QM^2Q', \tag{6}$$

and employing in (5) the new matrices of latent vectors and roots, $\underline{Q}$ and $\underline{M}^2$, given in (6).

In the oblimin class of solutions, the function

$$\xi = \sum_{j<k=1}^{q-1} \sum_{k=j+1}^{q} \left[\sum_{i=1}^{p} v_{ij}^2 v_{ik}^2 - (\gamma/p)\left(\sum_{i=1}^{p} v_{ij}^2 \right)\left(\sum_{i=1}^{p} v_{ik}^2 \right) \right] \tag{7}$$

is minimized. In its earliest form [Carroll, 1957], (7) had meaning with respect to the elements $\underline{v}$ of the reference-vector structure matrix, so that once this simplicity criterion was satisfied for the reference structure solution, rescaling of columns was necessary to obtain $\underline{P}$. Such solutions have been termed "indirect" oblimin sqlutions, the best-known cases of which have been those satisfying the quartimin ($\gamma = 0$), covarimin

(γ = 1), and biquartimin (γ = .5) criteria. Empirical evidence has suggested that quartimin solutions tend to be more oblique than desirable, and covarimin solutions tend to be too nearly orthogonal. Of the indirect oblimin criteria, biquartimin has been found to be close to optimal [Harman, 1967].

Jennrich and Sampson [1966] presented procedures for minimizing simplicity functions such as (7) in terms of the elements of the primary-factor pattern matrix, P, rather than of the reference structure matrix. Such criteria have been termed "direct" oblimin by Harman [1967]. Jennrich and Sampson developed their procedure with special reference to a "direct" quartimin criterion (γ = 0). Aside from the obvious advantage of providing solutions in which the simplicity function is optimized directly in terms of the matrix of major interest, the direct oblimin criterion, with γ set to 0 was shown to have two further important features: (1) it would result in a perfect unifactor configuration (zero factorial complexity) if such were possible for a set of data, and (2) it would not permit transformation to singularity, the latter condition certainly possible with indirect oblimin procedures and with direct oblimin criteria using positive values of γ.

In the present study, several direct oblimin solutions were obtained for most data sets; in nearly all cases, those with γ = 0 were superior to those based on other values of γ. Thus, the oblimin rationale is represented, in this paper, by the (1) biquartimin (indirect oblimin) and (2) direct oblimin with γ = 0 (henceforth referred to simply as "direct oblimin") criteria. In both cases, for all data analyzed, the variable vectors were first extended to unit length (normalized) for the purposes of the minimization, and later denormalized.

Extensions of Earlier Rationales

A major contribution of the paper by Harris and Kaiser [1964], was to show that a more basic referent, for transformation to an oblique simple structure, than the $\underline{A}$ (or $\underline{QM}$) matrix of (1) is the matrix $\underline{Q}$. This becomes more apparent if, using an Eckart-Young factorization of the middle four matrices on the right-hand side of (3) as

$$MT_1D_1T_2 = T_1^*D_1^*T_2^*, \tag{8}$$

we rewrite (3) (as did Harris and Kaiser) as

$$P = QT_1^*D_1^*T_2^*D_2, \tag{9}$$

which is the so-called "Case III" model. The problem represented in (9) appears, at first blush, amenable to easy solution, in that, orthogonal simple structure criteria could be employed to generate $\underline{T}_1^*$ and $\underline{T}_2^*$, and $\underline{D}_2$ is a function of $\underline{Q}$, $\underline{T}_1^*$, $\underline{D}_1^*$, and $\underline{T}_2^*$. Difficulties arise, however, in that no obvious principle exists to guide the choice of $\underline{D}_1^*$, and, in fact, $\underline{T}_1^*$, although orthonormal, is generally not solved for by optimization of an orthogonal simple structure criterion (although $\underline{T}_2^*$ may be). It will be recalled that the Case II model arose from (9) by setting $\underline{T}_1^*$ to $\underline{I}$, and useful practical procedures arose from solution of $\underline{D}_1^*$ (a function of $\underline{M}$) and $\underline{T}_2^*$ (by orthogonal rotation criteria).

In the present study, two issues were raised in connection with (9). First, does the Case II specialization of (9) typically result in a set of solutions in which the optimal oblique resolution for a set of data can be found? Secondly, does (9) have potential for generating new, practical oblique transformation methods? In an attempt to answer the first question, the authors generated 12 $\underline{QM}$ matrices, which were linearly transformable to known ideal primary-factor patterns, and for which these ideal primary factors had known intercorrelations. These data were used in

Abandoning the rather appealing logic of "oblique solutions by orthogonal transformations", the authors next embarked on an attempt to apply oblique transformational algorithms to $\underline{Q}$. If we represent a non-orthogonal transformation of $\underline{Q}$ as $\underline{U}$, $q \times q$, we have

$$P = QUD_U, \tag{10}$$

where the diagonal matrix $\underline{D}_U$, $q \times q$, is a function of $\underline{Q}$ and $\underline{U}$. More specifically, $\underline{QU}$ can be seen as an oblique transformation of $\underline{QM}$, by some oblique transformation matrix $\underline{B}$, $q \times q$, so that

$$\begin{aligned} QU &= QMB, \text{ or} \\ B &= M^{-1}U. \end{aligned} \tag{11}$$

Now, the covariance matrix of primary factors, $\underline{\Psi}$, $q \times q$, can be expressed as

$$\Psi = (B'B)^{-1}, \tag{12}$$

or, by substitution from (11), as

$$\Psi = U^{-1}M^2(U')^{-1}. \tag{13}$$

If we require the primary factors to have unit variances, we must standardize $\underline{\Psi}$ in (13) by

$$D_U = [\text{diag}(\Psi)]^{\frac{1}{2}}. \tag{14}$$

Thus, we have

$$\Phi = D_U^{-1}U^{-1}M^2(U')^{-1}D_U^{-1}, \tag{15}$$

and $\underline{P}$ is given by (10). Incidentally, the usual constraints on oblique transformation matrices, $\underline{H}$, $q \times q$, in general, viz,

$$\begin{aligned} &\text{diag}(H'H) = I, \text{ or} \\ &\text{diag}(H'H)^{-1} = I, \end{aligned} \tag{16}$$

the comparisons reported later in this paper, and a more detailed discussion of their generation is deferred until Section 3. The data sets varied in inherent factorial complexity and extent of departure from orthogonality among the factors, but in every case, the ideal or "true" factor hyperplanes were perfectly clear, or, more operationally, the nonsignificant factor pattern coefficients were exactly zero. It was found that in only two of the 12 data sets (the two in which a perfect unifactor configuration was possible), could the transformation to the ideal pattern be expressed as in (5) (with the independent cluster version solving the unifactor cases); (5) admittedly represents only a subset of possible Case II solutions, albeit the only subset for which practical procedures have been developed. Exclusive of those embodied in (5), no principles have yet been advanced, to the authors' knowledge, for the solution of $\underline{D}_3$ in (4). The foregoing results, then, suggested that attention could, perhaps, be profitably shifted from (4) or (5) to (9).

Using the logic of $\underline{Q}$ as a better referent for transformation than $\underline{QM}$, the authors first attempted to solve (9) sequentially as to the matrices $\underline{T}_1^*$, $\underline{D}_1^*$, and $\underline{T}_2^*$. One obvious approach was to begin with a $\underline{P}'\underline{P}$ proportional to $\underline{\Phi}$ solution, perform an Eckart-Young factorization on $\underline{M}^{\frac{1}{2}}\underline{T}_2\underline{D}_2$ in (5), obtaining $\underline{T}_1^*\underline{D}_1^*\underline{T}_2^*$, and then attempting to improve $\underline{T}_2^*$ by use of an orthogonal rotation criterion. The net effect, with this approach, was a negligible change in $\underline{T}_2^*$ and the transformed pattern. As mentioned earlier, no general principle was found for the solution of either $\underline{T}_1^*$ or $\underline{D}_1^*$. What appears to be true--in an "ideal" pattern that can be expressed as in (9)--is that $\underline{T}_1^*$ transforms the columns of $\underline{Q}$, not to a simple structure configuration, but rather to a position that can benefit maximally from the rescaling accomplished by $\underline{D}_1^*$, with this optimally prepared structure rotated into the ideal simple structure by means of $\underline{T}_2^*$.

need not apply to U, the matrix D_U accomplishing any column rescaling that is necessary.

The promax and oblimin (indirect) procedures were employed with Q on several artificial and also "real" data sets. The promax solutions obtained were not quite as good as standard promax solutions. The oblimin solutions were consistently good, but not appreciably better than either standard oblimin solutions or Harris-Kaiser Case II solutions.

The next line of attack consisted of attempting to combine certain features of existing techniques and the notion of obliquely transforming Q. Two procedures emerged from the experimentation, having resulted in consistently good solutions on all data sets employed. The procedures--designated Type A and Type B in the sequel--must be seen as attempts to render maximally clear, factors that are reasonably close to their mutually "optimal" positions. That is, in each case, it was assumed that primary factors located by the Harris-Kaiser independent cluster procedure--e.g., by a raw quartimax rotation of Q--are basically "correct" for the data at hand, but may be somewhat displaced from positions that would result in an overall optimally clear interpretation. An example of the independent cluster procedure yielding basically correct, but badly distorted, factors is Thurstone's [1947] well-known Box Problem--one of the data sets on which comparisons among the methods were made in this study.

Type A Solutions

In the Type A application, Q is first rotated to a raw quartimax (or varimax, etc.,) solution, QT, with elements $(qt)_{ij}$. A "pivot" point, p_j, is then located for each factor j, i.e., a value that appears to separate the significant from nonsignificant coefficients (in absolute value) of that factor. A constant that is roughly 1/4 to 1/3 the size of the largest (in absolute value) entry of the column in QT was found to

be generally best. A target matrix, $\underline{K}$, $\underline{p} \times \underline{q}$, is then established, in which each element of $\underline{K}$, $\underline{k}_{ij}$, is obtained by

$$|k_{ij}| = (|(qt)_{ij}|)^{p_j}/|(qt)_{ij}|. \tag{17}$$

The algebraic sign of $\underline{k}_{ij}$ is that of $(\underline{qt})_{ij}$. Use of (17) enhances the division, already present in $\underline{QT}$, between significant and nonsignificant coefficients, by lowering (in absolute value) those smaller than the pivot scalar and raising those that are larger. $\underline{Q}$ is then fitted, in a least squares sense, to $\underline{K}$, using the unrestricted least squares formulation

$$\begin{aligned} U &= (Q'Q)^{-1}Q'K \\ &= Q'K. \end{aligned} \tag{18}$$

The diagonal rescaling matrix, $\underline{D}_U$, is then obtained by (14), and $\underline{P}$, by (10). Although the authors have not been able to establish the nonsingularity of $\underline{U}$ in (18), it appears reasonable to assume that factor collapse would not occur with Type A solutions, given the basis for generation of $\underline{K}$. It should be noted that one could place constraints on $\underline{U}$, such as those associated with the Procrustes primary pattern application, viz., $\text{diag}[\underline{U}^{-1}\underline{M}^2(\underline{U}')^{-1}] = \underline{I}$. However, since the major consideration, in the present context, is not the preciseness of the least squares fit, but rather the approximation of an improved simple structure, the procedures outlined in (10) through (15) were deemed sufficiently precise.

The preceding steps may be employed more than once, in successively improving the solution. That is, after obtaining $\underline{U}$ by (18), a new target matrix can be generated in terms of $\underline{QU}$, as opposed to $\underline{QT}$. The user can thus repeat the procedure until no further improvement in overall quality of $\underline{P}$ is noticed. With several data sets analyzed in this study, the Type A procedure was applied more that once.

Type B Solutions

Type B solutions are more involved than the Type A, and employ a modification of a maximization algorithm developed elsewhere by the senior author [Hakstian, 1972]. In the latter paper, procedures were presented for maximizing the absolute distance between the mean square of predetermined significant and that of nonsignificant pattern coefficients. As with Type A, Type B solutions begin with an orthogonal rotation of $\underline{Q}$. Rather than inspecting the columns of $\underline{QT}$ to determine significant and nonsignificant coefficients for each factor, the significance of each variable for each transformed factor is taken to be proportional to the fourth power of its loading in $\underline{QT}$. The nonsignificance of each variable, $\underline{i}$, for each factor, $\underline{j}$ is taken as proportional to the fourth power of the [maximum loading--in absolute value--in that column, $|(\underline{qt})|_{j\max}$, - the variable's loading in absolute value, in the $\underline{j}$th column of $\underline{QT}$, $|(\underline{qt})|_{ij}$]. Thus, for transformed factor $\underline{j}$, for example, diagonal scaling matrices are established, $\underline{D}_{s_j}$ and $\underline{D}_{ns_j}$, both $\underline{p} \times \underline{p}$, as

$$D_{s_j} = \begin{bmatrix} (qt)_{1j}^4 & & & \\ & (qt)_{2j}^4 & & \\ & & \ddots & \\ & & & (qt)_{pj}^4 \end{bmatrix}, \text{ and}$$

$$D_{ns_j} = \begin{bmatrix} [|(qt)|_{jmax} - |(qt)|_{1j}]^4 & & & \\ & [|(qt)|_{jmax} - |(qt)|_{2j}]^4 & & \\ & & \ddots & \\ & & & [|(qt)|_{jmax} - |(qt)|_{pj}]^4 \end{bmatrix}.$$

The distinction between significant and nonsignificant coefficients for factor j is then enhanced by solving for the jth column of U, u_j, so as to maximize the function

$$\zeta = u_j' Q' D_{s_j}^2 Q u_j - w u_j' Q' D_{ns_j}^2 Q u_j, \tag{20}$$

subject to the constraint

$$u_j' u_j = 1, \tag{21}$$

to specify, uniquely, a solution vector u_j. The scalar w is a weight applied to the nonsignificance term in (20) to increase the relative importance of fitting factor j's hyperplane as closely as possible; large values of w (e.g., 100) appeared to yield best results with data. From well-known theory, maximization of (20) subject to (21) is attained by choosing for u_j the latent vector associated with the largest latent root of X_j, $q \times q$, where

$$\begin{aligned} X_j &= Q' D_{s_j}^2 Q - w Q' D_{ns_j}^2 Q \\ &= Q' (D_{s_j}^2 - w D_{ns_j}^2) Q. \end{aligned} \tag{22}$$

Thus, in Type B solutions, D_{s_j} and D_{ns_j} are determined for each factor, $j = 1, \ldots, q$. The matrices X_j are formed and the transformation vectors u_j obtained. We then have $U = [u_1, \ldots, u_q]$; D_U is obtained by

(14) and $\underline{P}$ by (10). As with Type A, the Type B algorithm may be applied several times until an optimal solution results. If a second Type B application were desired, the matrix $\underline{U}$ obtained from the first application would take the place of the initial orthonormal transformation, $\underline{T}$, and the matrices $\underline{D}_{s_j}$ and $\underline{D}_{ns_j}$ would contain entries $(\underline{qu})_{ij}$, rather than $(\underline{qt})_{ij}$. New $\underline{u}_j$ vectors, $\underline{u}^*_j$, would be generated by use of (22), and the procedure repeated until an optimal solution resulted. With several data sets in the present study, the Type B procedure was applied more than once. In the case of repeated application of both Type A and B procedures, the authors used total solution hyperplane-count (number of entries in the range $0 \pm .10$) as a rough indication of quality of solution. The procedures were repeated as long as the hyperplane-count increased, and terminated when this count leveled off or began to drop. Unless otherwise stated, however, in Section 4, Type A and B solutions were based on only one application.

Before leaving the discussion of procedures compared in the present study, one additional point should be noted. Particularly with the established methods, we are dealing, in each case, with a system or general paradigm for oblique transformation, rather than a single procedure. Since there are an enormous number of Harris-Kaiser (with a large number of possible $\underline{c}$ values), promax (with a large number of possible $\underline{b}$ values), and oblimin (with a large number of possible $\underline{\gamma}$ values) solutions possible, the present task was made manageable only by fixing on one or two specializations for each general method, that is, employing as representative of each general method, the specializations that are either those in common use or those that performed best with the data in the present study. These specializations have been already noted for each general method, and, strictly speaking, inferences stemming from the

comparisons should be confined to these specializations. Where possible, in Sections 4 and 5, generalizations to the larger paradigms are made.

3. Method of Comparison

The four established and two new procedures were compared by employing each with 17 sets of data, and establishing the closeness, for each method and data set, to an ideal simple structure resolution.

Data Used

Both artificial and "real" data sets were employed. Two groups of artificial data sets were constructed as follows. First, six "ideal" primary-factor patterns, $\underline{P}^*_i$, each 16 x 4, were constructed. These varied in terms of factorial complexity from Set 1, with a perfect unifactor (non-complex) structure, through Set 4, with some factorial complexity on eight of the variables, to Set 6, with the greatest degree of factorial complexity. Since all pattern coefficients were positive, the inherent factorial complexity was established for each set, i, by summing the off-diagonal elements of $\underline{P}^{*\prime}_i \underline{P}^*_i$. In all sets, the nonsignificant coefficients were exactly zero. Variable communalities ranged from about .20 to about .80, in each set, with an overall mean communality around .50. Next, two gramian primary-factor intercorrelation matrices $\underline{\Phi}^*_X$ and $\underline{\Phi}^*_Y$, both 4 x 4, were constructed both having only positive correlations. The mean $\underline{r}$ in $\underline{\Phi}^*_X$ was .40, in $\underline{\Phi}^*_Y$, .20. Next, 12 covariance matrices (for the "common parts" of the variables), $\underline{C}^*$, 16 x 16, of rank 4, were constructed by

$$C^*_{X_i} = P^*_i \Phi^*_X P^{*\prime}_i, \text{ and}$$

(23)

$$C^*_Y = P^*_i \Phi^*_Y P^{*\prime}_i$$

$$(i = 1, \ldots, 6).$$

Using $\underline{C}^*_{X_i}$ and $\underline{C}^*_{Y_i}$ ($\underline{i}$ = 1, ..., 6), 12 unrotated factor patterns, $\underline{F}^*_{X_i}$ and $\underline{F}^*_{Y_i}$, 16 x 4, were obtained by taking the four principal-axes factors of each of $\underline{C}^*_{X_i}$ and $\underline{C}^*_{Y_i}$, $\underline{i}$ = 1, ..., 6. Each $\underline{F}^*_{X_i}$ and $\underline{F}^*_{Y_i}$ was thus an exact linear transformation of the corresponding ideal primary pattern $\underline{P}^*_i$. The 12 matrices, $\underline{F}^*_{X_i}$ and $\underline{F}^*_{Y_i}$, $\underline{i}$ = 1, ..., 6, then, constituted the raw data with which the various procedures were applied. The "X" set of six matrices constituted quite oblique data, with the mean "true" factor intercorrelation .40. The "Y" set of six matrices constituted a more nearly orthogonal set, with the mean "true" factor intercorrelation .20.

In addition to the artificial data sets described above, five sets of actual data obtained from the literature were employed. These "real" data sets were chosen so as to vary in terms of (1) size of the data matrices--numbers of variables and factors--(2) reliability of the variables, and (3) apparent underlying factorial complexity of the variables. The data sets were: the Thurstone [1947] Twenty-Variable Box Problem (20 x 3), the Holzinger-Harman [see Harman, 1967] 24 Psychological Tests (24 x 4), a set of personality variables (20 x 5) first analyzed by Horn [1963]--referred to in the sequel as the 16 PF Data--a set of perceptual and ability measures (25 x 8) analyzed by Pemberton [1952]--referred to hereafter as the Pemberton Data--and a set of clinical measures (32 x 9) analyzed by Degan [1952]--referred to as the Degan Data. The first three data sets noted were used in the senior author's earlier study. For each real data set, a graphically transformed oblique solution was available for use as criterion. The unrotated (centroid) matrix for each set was found in the reference given for it above. A summary of the data sets used in the study appears in Table 1.

Insert Table 1 about here

Evaluation Procedures

For the artificial data sets, obvious "correct" solutions for the data were available, the matrices P^*_i. For the real data sets, the graphic primary pattern matrix was used as criterion. For each data set, two Harris-Kaiser--independent cluster (IC) and $P'P$ proportional to Φ $(P'P)$--two promax (b = 2 and 4), biquartimin, direct oblimin, Type A, and Type B solutions were obtained. The overall evaluation criterion employed in the study was closeness to an ideal simple structure [Thurstone, 1947]. This overall criterion was operationalized as:

(1) apparent "cleanness" of the solution, as indicated by total solution hyperplane-count--or number of primary pattern coefficients in the range 0 ± .10, and

(2) congruence of the solution to a known "correct" solution, or ideal simple structure, for the data. This latter criterion was implemented as follows:

(a) The degree of departure from orthogonality of the solution was compared to that of the "correct" or criterion solution. The interfactor correlations were converted to angular separations (θ = arccos r); these were then subtracted from 90°, converted to absolute values (in the case of negative correlations), and averaged. Thus, mean angular departure from orthogonality of the factors in each solution was examined, and compared with that of the criterion solution.

(b) The overall closeness--as measured by the mean angular separation from criterion--of the solution to the "correct" or criterion solution was assessed. That is, in each obtained solution, the factors were matched with the most closely corresponding factors in the criterion solution. Since both the obtained and criterion solutions were linear transformations of the same initial unrotated matrix, in each case, true correlations--and angular

separations--between corresponding factors could be obtained. The q between-solution angular separations were averaged, and the index noted above was thus obtained. This latter criterion, of which a more extensive discussion can be found in the senior author's earlier paper [Hakstian, 1971], was generally given greatest weight in the present study.

4. Results and Discussion

The results of the present study are presented, in summary form, in Tables 2 through 4. The results for the artificial and real data sets are treated separately.

Artificial Data Sets

The results for the most oblique (X) data sets are displayed in Table 2; those for the more nearly orthogonal (Y) sets are given in Table 3. For the Type A solutions presented, the p_j values in (17) were set to .10 throughout.

Insert Tables 2 and 3 about here

If we begin with the most oblique data sets, it can be seen, from Table 2, that considerable discrimination between methods is possible. Regarding the hyperplane-counts, even with the simplest data set (1X), in which a true unifactor configuration exists, some methods (Harris-Kaiser - $P'P$, and biquartimin) were unable to get an accurate fix on all the hyperplanes. This tendency is accentuated with the more complex sets.

With regard to the obliquity of the solutions obtained, it is apparent, from Table 2, that large differences exist among the methods compared. The earlier finding [Hakstian, 1971] that, with promax, the size of b is directly related to departure from orthogonality is supported. Biquartimin can be seen to yield consistently and substantially ortho-

gonally biased solutions with these data sets. It is interesting to note that with the Harris-Kaiser IC solutions, the departure from factor orthogonality appears to increase as a function of the inherent factorial complexity of the pattern matrix. This relationship is not perfect, but there is a trend, with this procedure, in the direction of locating pairs of factors between their respective defining clusters, when other variables are to be found in these between-cluster locations. This relationship between obliquity and complexity does not appear to exist, to the same extent, with the other methods compared, although the biquartimin and P'P solutions exhibit the same relationship to a lesser extent. The P'P version of the Harris-Kaiser procedure can be seen, in general, to have an orthogonal bias, as can be expected from its formulation. Overall, only the direct oblimin and Type A and B solutions appear to exhibit essentially correct obliquity throughout the six data sets.

In terms of closeness of the factors in the obtained solutions to those in the "correct" solution, again, as is seen from Table 2, considerable variability among the methods is apparent. As found earlier [Hakstian, 1971] with the two Harris-Kaiser procedures, as complexity increases, the P'P solution becomes the better of the two, whereas for factorially simple data, the IC solution is better. Of the four established techniques, direct oblimin yielded solutions closest to the "correct" factors for these six most oblique artificial sets, followed by promax, Harris-Kaiser, and lastly biquartimin. The Type A and B solutions were substantially closer to the "correct" solution than were even the direct oblimin.

Over the three criteria employed with this first set of artificial data, the direct oblimin solutions tended to be the best of those obtained

with the established methods. The promax solutions were generally next best, followed by the Harris-Kaiser and biquartimin solutions. The Type A and B solutions tended to be best overall.

For the more nearly orthogonal artificial data sets, much the same picture is seen to emerge as with the more oblique sets. Less discrimination was possible between methods on the hyperplane-count criterion, most solutions having the correct number of hyperplane entries.

The findings regarding angular departure from orthogonality noted earlier can be seen to hold for these data sets as well. The Harris-Kaiser $\underline{P}'\underline{P}$, and more notably biquartimin solutions consistently exhibited an orthogonal bias. The Harris-Kaiser IC solutions exhibited the earlier noted relationship between obliquity and complexity, again to a marked degree, and the same relationship is noted--to a less degree--with the $\underline{P}'\underline{P}$, promax - $\underline{b}$ = 4, and biquartimin solutions.

Regarding closeness of the obtained factors to "correct" factors, the same overall ranking, in terms of established procedures, was obtained with the more nearly orthogonal data: direct oblimin, promax, Harris-Kaiser, biquartimin. If we consider promax as a single method, however, and choose the best version ($\underline{b}$ = 2 or 4) for each data set, the overall (over the six data sets) closeness to the criterion is slightly better for promax than for direct oblimin. The Type A and B solutions were again substantially closer to the "correct" solutions than were the direct oblimin. It may be noted that for three of the four established procedures, solutions for the more nearly orthogonal data sets were, in general, closer to the "correct" solutions than they were in the more oblique sets. Only the direct oblimin solutions were closer to the "correct" solutions when the factors were more highly correlated. The Type A and B solutions were of roughly equal closeness to the criteria in the two cases.

In summary, the results for the artificial data sets appear to suggest two facts: (1) for the established methods, over the three criteria employed, direct oblimin appears to yield the best solutions, followed closely by promax, which in turn is followed by Harris-Kaiser, and lastly biquartimin, and (2) substantial improvement in solution over the existing techniques can be obtained by the Type A and B modifications. These conclusions must be held with considerable caution, however, because of the nature of the artificial data sets. The fact, for example, that the hyperplanes were absolutely cleanly defined renders any results based on such data as enlightening, but certainly not definitive. We wonder, in connection with this fact, whether such exact zero hyperplanes give some advantage to procedures that work with between-factor simplicity, such as quartimin (represented in "direct" form by the direct oblimin criterion in this study).

Real Data Sets

The results for the five real data sets are displayed in Table 4. For the Type A solutions, the p_j values in (17) were: Box Problem, .10, 24 Psychological Tests, .10, 16 PF Data, .20, Pemberton Data, .15, and Degan Data, .20.

Insert Table 4 about here

In terms of hyperplane-counts, considerable variability among methods was again noted. This phenomenon is most apparent in the Box Problem, where the counts ranged from five to 27. With the real data sets, unlike with the artificial, the criterion (graphic) solution did not always have the largest hyperplane-count, although in only one set (24 Psychological Tests) was this not the case. Hyperplane-count must be regarded as an only imperfect indicator of overall quality of solution.

The trend noted earlier, with the Harris-Kaiser solutions, of the IC procedure locating factors between defining clusters, and the P'P procedure locating them outside (producing, respectively, over-oblique and over-orthogonal bias) appears to hold with the real sets, although with the Box Problem, slight oblique bias is noted for the P'P solution, and with the 16 PF Data, slight orthogonal bias, for the IC solution. This trend perhaps suggests that, in general, a better value for the c parameter in (5) might be a number between 0 and .5, such as .25. With all real data sets except the Box Problem, both oblimin procedures--biquartimin and direct oblimin--appear to have an orthogonal bias. As before, the Type A and B solutions were, in general, close to correct obliquity, although both tended to be too oblique with the Pemberton Data.

The relative closeness of the various solutions to the criterion solution tended to differ, with the real data, from that with the artificial. It can be seen from Table 4 that, of the established methods, one of the two Harris-Kaiser solutions was closest overall to the criterion solution for each of the five data sets. The promax and direct oblimin solutions tended to be next closest overall, with the promax superior on the Box Problem and 24 Psychological Tests, and direct oblimin on the 16 PF, Pemberton, and Degan Data, although for the last two data sets, the difference in closeness of solution to the criterion is very slight. The biquartimin solutions tended to be furthest displaced from the criterion, although with the Box Problem, a good biquartimin solution resulted. It is noted that no biquartimin solution was obtained for the Pemberton Data, because the algorithm produced factor collapse with these data. Again, the Type A and B solutions were closer to the criterion solution than were those of the established methods. This was accomplished in some cases by

applying these procedures several times in succession and, thus, gradually improving the solution until no further improvement was possible. Type A and B solutions considerably closer to the criterion than those yielded by the established methods were obtained particularly with the 16 PF and Degan Data.

It appears true, in summary, with the real data sets, that over the three criteria, the Harris-Kaiser procedure resulted in the best overall solutions, of those obtained by the established methods. The promax and direct oblimin solutions tended to rank next, followed by the biquartimin solutions. The Type A and B solutions were, generally, best of all. Although based on quite different rationales, the two new procedures tended to yield fairly similar solutions. Overall, with both the artificial and real data sets, the two procedures yielded approximately equally good solutions.

The discrepancy in overall ranking for the direct oblimin solutions--between the artificial and real data sets--prompted the authors to perform some additional analyses, with the latter data, employing values of γ, in (7), other than zero. The objective here was to see whether the direct oblimin paradigm was capable of yielding better solutions than those with γ set to zero. The three data sets for which the direct oblimin function with $\gamma = 0$ yielded the relatively poorest solutions were the Box Problem, 24 Psychological Tests, and Degan Data. For the Box Problem, the overobliquity of the $\gamma = 0$ solution was partially corrected by employing negative values for γ--beginning with -1.0, and proceeding in steps of 1.0 down to -7.0. The best solution--although still somewhat too oblique--was obtained with $\gamma = -2.0$; the mean angular separation from criterion was $3^{\circ}15'$, a separation almost as small as those obtained by the Harris-Kaiser and biquartimin procedures. For the 24 Psychological Tests, greater

obliquity than that obtained with $\gamma = 0$ was needed. The best direct oblimin solution for these data was that with $\gamma = .3$. With this solution, the obliquity was essentially correct, and the mean angular separation from criterion was 2°44'--representing the best overall solution for these data. For the Degan Data, again greater obliquity was required, and values for γ of .1, .2, .3, .4, and .5 were tested. As γ was thus increased, so was the overall obliquity of solution, until at $\gamma = .5$, the solution collapsed. No overall improvement in solution, for these data, over that obtained with $\gamma = 0$ was realized, however, in terms of closeness to the criterion solution.

It appears, therefore, that the user employing the direct oblimin criterion may need to obtain solutions with γ varied between negative and positive values, to ensure getting the best direct oblimin solution for his data, although the value for γ of zero may still be best in general. The problem of factor collapse is present, as has been seen, with positive values much greater than zero.

An additional point worth noting in this section has to do with the behavior of the oblimin criterion functions during iteration. First, as was seen in Table 4, the danger of transformation to singularity--or factor collapse--is a very real one with the indirect oblimin criterion, represented in the present study by biquartimin (and, as was noted above, for the direct oblimin criterion utilizing positive values for γ). This possibility must be seen as a limitation of biquartimin, although the authors have no idea of the proportion of actual data sets for which collapse can be expected to occur. As noted earlier, with the direct oblimin criterion in which γ, in (7), is set to 0, collapse cannot occur, nor can it with the Harris-Kaiser procedure.

Secondly, the authors studied the behavior of the direct oblimin

criterion function in the solutions obtained. Jennrich and Sampson [1966] noted that no claims regarding guaranteed convergence could be made for the direct oblimin function (just as none can be made for virtually any analytic oblique criterion). In the present study, with the convergence criterion equal to .00001--an admittedly small number, but the one set as the default value in the computer program used--in only two of the 12 artificial data sets (1X and 4X) and one real set (24 Psychological Tests) was convergence obtained. In the other data sets, the program was allowed to run to 99 iterations before termination, and in each case, the criterion function, although not fully converged, was changing by only between .001 and .0001. It was thus assumed that the solutions obtained were essentially what they would be at convergence. To further examine the issue of convergence, the iterations were restarted with the Box Problem, and a change of only .014 between the 99th and 200th iterations was noted in the (still unconverged) criterion function, with virtually no change in the corresponding solutions. We are still not clear, therefore, on the behavior of this function, i.e., whether it eventually tends to converge, given enough time, or, perhaps, reaches a point of divergence or oscillation.

5. Conclusions

One point that should be reiterated in connection with the results of the present study is that with both the Harris-Kaiser and promax methods, only varimax orthogonal rotations were employed for generating the T_2 matrix in (5), for the Harris-Kaiser procedure, and for establishing the preliminary solution, for promax. In the senior author's earlier study [Hakstian, 1971], there appeared to be some evidence that a "transvarimax" criterion, such as equamax, might be superior to varimax for these functions. Subsequent unpublished research, however, has revealed that, with the Harris-Kaiser procedure, the effects of orthogonal rotation criterion may be less

than first thought, with the value for c in (5) vastly more important. The issue of intermediate orthogonal rotation criterion, in connection with both the Harris-Kaiser and pro. ax methods, is in need of careful further investigation.

The following general conclusions appear warranted from the results of the present study:

The Methods Compared

1. The Harris-Kaiser "Case II" specialization of (5) represents a relatively dependable paradigm for oblique transformation, with the benefits of rapid convergence to solution and no possibility of transformation to singularity. Although the two widely-used special cases--IC and P'P--may not be optimal for a given data set, it is likely true that some value for c, in (5), between 0 and .75 will yield a good solution for the data. There certainly is evidence that the IC version will often not be optimal, as it appears to have a consistent over-oblique bias. The P'P version, on the other hand, will tend to be orthogonally biased, and if the specialized "Case II" formulation represented in (5) is used, a generally better value for c (that is, if only one value could be employed) is likely to be around .25. In general, c should vary directly as the inherent factorial complexity of the data (as nearly as this can be determined). Another promising line of further inquiry would be based on the notion of retaining Q as the optimal basic referent for oblique transformation, and pursuing solutions for T_1^*, D_1^*, and T_2^* in (9), i.e., developing the "Case III" model.

2. The promax procedure again appears to be relatively effective with data, despite its rather simple rationale. The values for b of two and four are, arbitrary, of course, and undoubtedly not optimal, only approximately correct. Obtaining several promax solutions, with b varied between two and, say, six, would appear to be a good strategy for the user.

In fact, it is likely true that a different b value for each factor (or even each element, as is accomplished in the Type A procedure with QT) would be superior to a common value for all factors. It appears unlikely that factor collapse could occur with the promax procedure, although we cannot prove this. Further research could profitably be directed at (1) methods for generating the preliminary solution (it need not be orthogonal) and (2) better procedures for establishing the target matrix from this solution.

3. The biquartimin procedure--assumed to be the generally best indirect oblimin method--appears to be rather ineffective. The danger of factor collapse is present, and the consistent orthogonal bias appears to preclude optimal oblique solutions.

4. The direct oblimin class of criteria appears capable of yielding good solutions for most kinds of data. The criterion with $\gamma = 0$, in (7), is probably the best overall version of this general class, with the absence of the possibility of factor collapse and a not overly-great bias regarding obliquity of factors (although a degree of orthogonal bias appeared with the real data sets in the present study). As pointed out by Harman [1967], with negative values for γ, solutions become more nearly orthogonal, until with large negative values (less than, say, -10), solutions begin to return to their unrotated form. The γ parameter may be allowed to be positive (in small fractional steps, e.g., .1, ..., .6), in which case obliquity increases, with, for all practical purposes, factor collapse inevitable as γ increases. With the Degan Data in the present study, for example, practical collapse occurred at $\gamma = .5$.

5. The notion of transforming Q obliquely--thus obtaining solutions that belong to Harris and Kaiser's [1964] "Case III" set--and that of

successively improving obtained oblique solutions appear worthy of attention. Two procedures employing these notions--the Type A and B methods examined in this study--yielded consistently good solutions. It appears that the configuration represented by QT can be counted upon to give good initial approximations to the eventual optimal placement of axes. Further research is obviously needed on the best means for successively altering QT.

Overall Conclusions

It is probably best to attempt to assess the previous results from two points of view: (1) in terms of the specific forms of the various criteria compared, and (2) in terms of the overall paradigms for transformation represented by these specific cases. With regard to the first frame of reference, the authors had some difficulty in reconciling the findings based on the artificial data sets with those based on the real. In the final analysis, more weight was given the real data sets, as the inherent clarity of the hyperplanes in the artificial sets was considered to limit their generalizability. With this fact in mind, the authors feel that--of the specific cases examined of the established methods--the Harris-Kaiser procedure (IC and P'P) must again be viewed as, overall, best, with the direct oblimin ($\gamma = 0$) and promax ($b = 2$ and 4) procedures ranking next. Biquartimin appears to be the least effective procedure studied, and the oblimin logic appears best represented in the direct oblimin formulation.

One conclusion that appears inescapable is that no single computing procedure--i.e., specialization of a general paradigm--of those examined in this study, can be counted on to yield uniformly optimal oblique solutions for all kinds of data. There appear to be extremely complicated interactions between the specification parameter, in each case [c in (5), γ in (7), and b in the promax method], and various characteristics of the data. Thus, if one were to obtain only one solution for a given paradigm--perhaps

with c = .25 for the Harris-Kaiser procedure, b = 2 or 3 for promax, and γ = 0 for direct oblimin--he would run considerable risk, as is seen from the present results, of having a less than optimal solution for his data.

Although the method of comparison employed in the present study precludes rigorous generalizations regarding the overall paradigms of which special cases were compared, it appears true that the general Harris-Kaiser, promax, and direct oblimin formulations are each well enough conceived to include a specialization that yields a good oblique solution for almost any given set of data. The user would be wise, in this regard, to obtain several solutions for a paradigm: for the Harris-Kaiser procedure, with c = 0, .25, .5, .75, perhaps; for promax, with b = 2, ..., 6; for direct oblimin, with γ = -2, 0, .1, .2, ..., .6. Possible drawbacks of the direct oblimin criterion--over many γ values--would be the possibility of factor collapse for positive values, and the unknown convergence tendencies noted earlier. The authors feel that the Harris-Kaiser procedure, particularly with the greater possibility of future refinements in its formulation, is, perhaps, the most promising overall method.

The problem of choosing among several Harris-Kaiser, promax, or direct oblimin solutions remains, of course, and the general user has only imperfect means by which to make the choice. Criteria that could be employed include (1) clarity and extent of definition of the obtained hyperplanes, (2) (lack of) factorial complexity of the variables, (3) (subjective) overall interpretability of the obtained solution. Another criterion somewhat related to (3), that may have merit--<u>all other things being equal</u>--is low overall factor obliquity; <u>i.e.</u>, of two solutions exhibiting equally good simple structures, that with lower overall factor intercorrelations may be preferable, in terms of factor interpretability. If higher-order analyses are planned, however, this last point is probably

inapplicable, and the correct choice among solutions--in terms of quality of simple structure--is of utmost importance.

The Type A and B procedures yielded encouraging results in the present study, and tended to suggest that at least some of the bewildering complexity stemming from the interactions between the specification parameters and features of the data may be lessened by gradually improving a (perhaps initially poor) solution resulting from one particular procedure. Better solutions were generally obtained, in the present study, with this approach than yielded by the best versions of the paradigms compared. Additional methods by which gradual approximations towards optimal final solutions can be made--in terms of Q--are currently being investigated.

REFERENCES

Carroll, J. B. Biquartimin criterion for rotation to oblique simple structure in factor analysis. Science, 1957, 126, 1114-1115.

Cattell, R. B., & Muerle, J. L. The "maxplane" program for factor rotation to oblique simple structure. Educational and Psychological Measurement, 1960, 20, 569-590.

Degan, J. W. Dimensions of functional psychosis. Psychometric Monographs, No. 6, 1952.

Guttman, L. Image theory for the structure of quantitative variates. Psychometrika, 1953, 18, 277-296.

Hakstian, A. R. A comparative evaluation of several prominent methods of oblique factor transformation. Psychometrika, 1971, 36, 175-193.

Hakstian, A. R. Optimizing the resolution between salient and non-salient factor pattern coefficients. British Journal of Mathematical and Statistical Psychology, 1972, 25, 229-245.

Harman, H. H. Modern factor analysis. (2nd ed.) Chicago: University of Chicago Press, 1967.

Harris, C. W. Some Rao-Guttman relationships. Psychometrika, 1962, 27, 247-263.

Harris, C. W., & Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 1964, 29, 347-362.

Hendrickson, A. E., & White, P. O. PROMAX: A quick method for rotation to oblique simple structure. British Journal of Statistical Psychology, 1964, 17, 65-70.

Horn, J. L. Second-order factors in questionnaire data. Educational and Psychological Measurement, 1963, 23, 117-134.

Jennrich, R. I., & Sampson, P. F. Rotation for simple loadings. Psychometrika, 1966, 31, 313-323.

Jöreskog, K. G., & Goldberger, A. S. Factor analysis by generalized least squares. Psychometrika, 1972, 37, 243-260.

Pemberton, C. The closure factors related to other cognitive processes. Psychometrika, 1952, 17, 267-288.

Saunders, D. R. The rationale for an "oblimax" method of transformation in factor analysis. Psychometrika, 1961, 26, 317-324.

Thurstone, L. L. Multiple-factor analysis. Chicago: University of Chicago Press, 1947.

TABLE 1

Summary of Data Sets Used in the Study

	Artificial Sets		
	Most Oblique (X)		Least Oblique (Y)
Least Factorially Complex ↓ Most Factorially Complex	1X		1Y
	2X		2Y
	3X	(all 16x4)	3Y
	4X		4Y
	5X		5Y
	6X		6Y

Real Sets

Box Problem (20x3)
24 Psychological Tests (24x4)
16 PF (20x5)
Pemberton (25x8)
Degan (32x9)

TABLE 2

Results of Analyses Involving the Most Oblique (X) Artificial Data Sets

Total-Solution Hyperplane-Counts

Solution	Data Set* 1X	2X	3X	4X	5X	6X
"Correct"	48	44	40	44	41	40
Harris-Kaiser (IC)	48	44	40	44	36	35
Harris-Kaiser (P'P)	46	41	39	44	40	34
Promax - b=2	48	44	40	44	36	40
Promax - b=4	48	44	40	44	35	40
Biquartimin	38	35	40	41	39	37
Direct Oblimin	48	44	40	44	41	40
Type A	48	44	40	44	41	40
Type B	48	44	40	44	41	40

Mean Angular Departures from Orthogonality

Solution	Data Set* 1X	2X	3X	4X	5X	6X
"Correct"	23°43'	23°43'	23°43'	23°43'	23°43'	23°43'
Harris-Kaiser (IC)	23°43'	28°11'	31° 8'	29°24'	35°55'	34° 8'
Harris-Kaiser (P'P)	10°57'	13° 4'	14°19'	13°39'	16°29'	15°45'
Promax - b=2	17°31'	18°36'	19°50'	17°59'	17° 7'	18°44'
Promax - b=4	23°24'	26°20'	29°39'	26°51'	26°52'	29°22'
Biquartimin	9°23'	11°52'	14° 9'	12°45'	16°20'	15°52'
Direct Oblimin	23°43'	24°16'	25°46'	23°11'	23°13'	24°35'
Type A	23°43'	23° 5'	23°30'	24° 1'	23° 7'	23° 0'
Type B	23°43'	24°14'	25°41'	23°43'	24° 8'	24°33'

Mean Angular Separations from Criterion Solution

Solution	Data Set* 1X	2X	3X	4X	5X	6X
Harris-Kaiser (IC)	0° 0'	4°20'	5°43'	4°21'	8°54'	8° 2'
Harris-Kaiser (P'P)	9°37'	8°15'	7°48'	7°30'	6°21'	7°21'
Promax - b=2	4°46'	4°33'	4°39'	5°29'	6°42'	5°13'
Promax - b=4	0°12'	3°13'	5° 1'	2°37'	8°28'	5°41'
Biquartimin	11°19'	9°39'	8°19'	8°49'	6°35'	7°12'
Direct Oblimin	0° 0'	1°12'	2°20'	0°36'	2°34'	2°42'
Type A	0° 0'	0°32'	0°33'	0°28'	1° 4'	1° 0'
Type B	0° 0'	0°45'	1°42'	0° 5'	0°21'	0°53'

*The data sets run from least factorially complex (1X) to most complex (6X).

TABLE 3

Results of Analyses Involving the More Nearly Orthogonal (Y) Artifical Data Sets

Total-Solution Hyperplane-Counts

Solution	Data Set* 1Y	2Y	3Y	4Y	5Y	6Y
"Correct"	48	44	40	44	41	40
Harris-Kaiser (IC)	48	44	40	44	38	36
Harris-Kaiser (P'P)	48	44	40	44	41	40
Promax - b=2	48	44	40	44	41	40
Promax - b=4	48	44	40	44	35	38
Biquartimin	48	44	40	44	41	40
Direct Oblimin	48	44	40	44	41	40
Type A	48	44	40	44	41	40
Type B	48	44	40	44	41	40

Mean Angular Departures from Orthogonality

Solution	Data Set* 1Y	2Y	3Y	4Y	5Y	6Y
"Correct"	12° 0'	12° 0'	12° 0'	12° 0'	12° 0'	12° 0'
Harris Kaiser (IC)	11°56'	15°49'	18°32'	17° 0'	23°21'	21°27'
Harris-Kaiser (P'P)	5°29'	7°25'	8°32'	7°58'	10°32'	9°51'
Promax - b=2	9°53'	11°33'	13°21'	11°23'	11°39'	12°52'
Promax - b=4	11°47'	15°20'	18°26'	16°39'	18° 6'	19°11'
Biquartimin	3°39'	5°28'	7° 2'	6°14'	9° 8'	8°28'
Direct Oblimin	11°54'	13°21'	15°14'	12°55'	14° 1'	14°54'
Type A	12° 0'	11°47'	11°24'	11°12'	11° 7'	11° 1'
Type B	12° 0'	12°11'	13°21'	11°35'	11°53'	12°22'

Mean Angular Separations from Criterion Solution

Solution	Data Set* 1Y	2Y	3Y	4Y	5Y	6Y
Harris-Kaiser (IC)	0° 1'	4°45'	5°57'	4°38'	9°23'	8°27'
Harris-Kaiser (P'P)	5°22'	4°38'	4°36'	3°27'	3°57'	4°56'
Promax - b=2	1°34'	2°16'	3°17'	1°15'	4°22'	3°40'
Promax - b=4	0° 3'	4°19'	6° 3'	4°28'	9°33'	7°28'
Biquartimin	7°20'	6°20'	5°44'	5°19'	4°24'	5°10'
Direct Oblimin	0° 1'	2°35'	3°44'	1°34'	3°40'	4°18'
Type A	0° 0'	0°32'	0°34'	0°34'	1° 9'	1° 7'
Type B	0° 0'	0°49'	1°46'	0° 6'	0°22'	0°56'

*The data sets run from least factorially complex (1Y) to most complex (6Y).

TABLE 4

Results of Analyses Involving Real Data Sets

Total-Solution Hyperplane-Counts

Solution	Data Set				
	Box Prob.	24 P. Tests	16PF	Pemberton	Degan
Graphic	27	42	48	129	182
Harris-Kaiser (IC)	5	37	45	104	140
Harris-Kaiser (P'P)	27	39	43	107	149
Promax - b=2	27	41	45	115	116
Promax - b=4	9	37	42	105	125
Biquartimin	27	44	40	**	141
Direct Oblimin	27	42	45	107	143
Type A*	27	39	45	119	160
Type B†	27	39	48	118	164

Mean Angular Departures from Orthogonality

Solution	Data Set				
	Box Prob.	24 P. Tests	16PF	Pemberton	Degan
Graphic	10°42'	32°18'	10° 9'	14°12'	13°48'
Harris-Kaiser (IC)	27°29'	34°16'	9°52'	24°14'	14°58'
Harris-Kaiser (P'P)	13°15'	15°39'	5°21'	8°41'	6°33'
Promax - b=2	14°58'	17°22'	6°13'	9° 9'	11°19'
Promax - b=4	25°40'	30°24'	10°59'	18°29'	11°51'
Biquartimin	13°18'	18°28'	7° 0'	**	11°23'
Direct Oblimin	17°53'	22° 5'	6°21'	9° 6'	7°47'
Type A*	11°33'	32°37'	11°37'	22°19'	15°25'
Type B†	12°21'	33°17'	11°41'	23°34'	14°13'

Mean Angular Separations from Criterion Solution

Solution	Data Set				
	Box Prob.	24 P. Tests	16PF	Pemberton	Degan
Harris-Kaiser (IC)	10°51'	4° 3'	22° 5'	24°25'	16°36'
Harris-Kaiser (P'P)	3° 0'	11°53'	13°22'	25°29'	19°44'
Promax b=2	3°32'	10°13'	18°20'	25°46'	22° 6'
Promax b=4	9°42'	5° 9'	19°41'	25°39'	20°54'
Biquartimin	3°10'	10°18'	19°50'	**	22°31'
Direct Oblimin	5° 0'	7°30'	13°56'	24°44'	20°28'
Type A*	2°21'	3°50'	8°41'	23°39'	8°46'
Type B†	2°11'	3°49'	9°44'	24° 7'	9°59'

*The Type A procedure was applied once for the 24 Psychological Tests and Pemberton Data, twice for the Box Problem, three times for the 16PF Data, and four times for the Degan Data.

†The Type B procedure was applied once for the 24 Psychological Tests and Pemberton Data, and twice for the Box Problem, 16PF Data, and Degan Data.

**Factor collapse occurred during the minimization.